In [1]:

```python
import pandas as pd
df = pd.read_csv('F:\\Mtech\\Papers\\zander\\2020_IN_Region_Mobility_Report.csv')
```

In [2]:

```python
# Convert the date column to datetime type
df.date = pd.to_datetime(df.date, format='%d-%m-%Y')
```

In [3]:

```python
df.sample(100)
```

Out[3]:

	country_region_code	country_region	state	date	retail	grocery	parks	transit	workplaces	residential
30390	IN	India	Chhattisgarh	2020-10-11	-38.0	-39.0	-76.0	-9.0	-17.0	16.0
40090	IN	India	Delhi	2020-06-11	-58.0	-14.0	-91.0	-66.0	-56.0	23.0
107885	IN	India	Maharashtra	2020-03-04	1.0	10.0	5.0	3.0	9.0	-1.0
12237	IN	India	Assam	2020-04-14	-67.0	-54.0	-11.0	-60.0	-58.0	23.0
47770	IN	India	Gujarat	2020-05-18	NaN	NaN	-43.0	-42.0	-13.0	NaN
...	...	...	...	...	...	...	...	...	...	...
168963	IN	India	Tamil Nadu	2020-03-05	8.0	8.0	5.0	9.0	1.0	-1.0
50688	IN	India	Gujarat	2020-07-11	-55.0	23.0	-77.0	-31.0	-15.0	19.0
155814	IN	India	Rajasthan	2020-12-13	-9.0	30.0	-16.0	-9.0	-3.0	12.0
19854	IN	India	Bihar	2020-12-04	-3.0	32.0	18.0	0.0	-11.0	10.0
17331	IN	India	Bihar	2020-03-03	2.0	-1.0	18.0	11.0	2.0	-4.0

100 rows × 10 columns

In [27]:

```python
covid_df = pd.read_csv('F:\casess.csv')
covid_df.sample(100)
```

Out[27]:

	date	state	cases	deaths
11826	18-02-2021	Kerala	1016848	4032
11397	06-02-2021	Madhya Pradesh	256017	3819
9914	27-12-2020	Himachal Pradesh	54280	913
6604	24-09-2020	Delhi	256789	5087
4706	01-08-2020	Andaman and Nicobar Islands	548	5
...	...	...	...	...
14767	11-05-2021	Chandigarh	51070	575
14914	15-05-2021	Delhi	1380981	20907
9771	23-12-2020	Jammu and Kashmir	118803	1850
7815	28-10-2020	Tamil Nadu	714235	10983
1296	25-04-2020	Haryana	272	3

100 rows × 4 columns

In [23]:

```python
# Convert the date column to datetime type
covid_df.date = pd.to_datetime(covid_df.date, format='%d-%m-%Y')
```

In [24]:

```python
final_df = df.groupby(['state', 'date']).mean().reset_index()
```

In [25]:

```python
final_df.head()
```

Out[25]:

	state	date	retail	grocery	parks	transit	workplaces	residential
0	Andaman and Nicobar Islands	2020-02-15	1.0	-79.0	3.0	-2.0	4.0	2.0
1	Andaman and Nicobar Islands	2020-02-16	0.0	-79.0	-1.0	3.0	1.0	2.0
2	Andaman and Nicobar Islands	2020-02-17	-2.0	-79.0	1.0	1.0	7.0	2.0
3	Andaman and Nicobar Islands	2020-02-18	-4.0	-79.0	-6.0	-3.0	3.0	2.0
4	Andaman and Nicobar Islands	2020-02-19	-7.0	-79.0	-6.0	1.0	3.0	2.0

In [26]:

```python
final_df = final_df.merge(covid_df, how="inner", on=['state', 'date'])
```

In [20]:

```python
final_df.tail()
```

Out[20]:

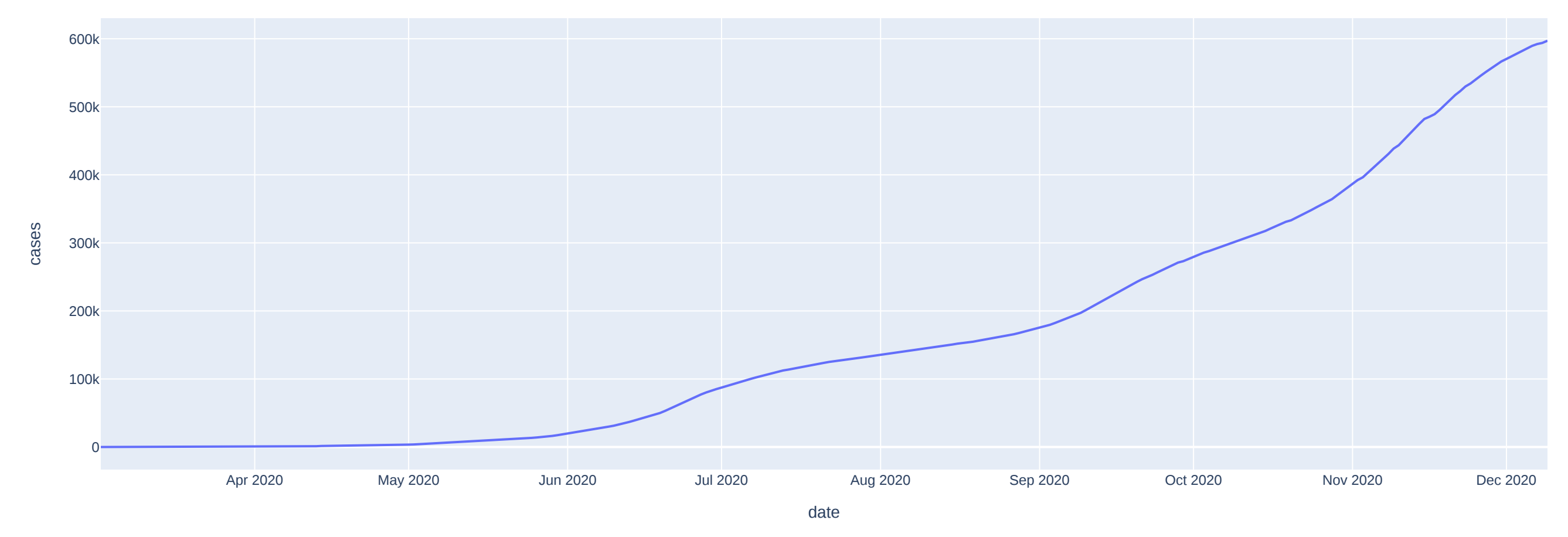
	state	date	retail	grocery	parks	transit	workplaces	residential
11526	West Bengal	2020-12-27	-21.095238	9.619048	7.714286	-3.047619	4.047619	11.619048
11527	West Bengal	2020-12-28	-20.000000	13.714286	7.285714	-6.000000	-10.190476	13.904762
11528	West Bengal	2020-12-29	-21.380952	9.428571	6.095238	-7.523810	-11.190476	14.523810
11529	West Bengal	2020-12-30	-18.904762	14.857143	3.857143	-0.571429	-2.428571	12.666667
11530	West Bengal	2020-12-31	-12.714286	23.904762	10.047619	4.142857	-0.714286	12.571429

In [10]:

```python
final_df.to_pickle('app.data')
```

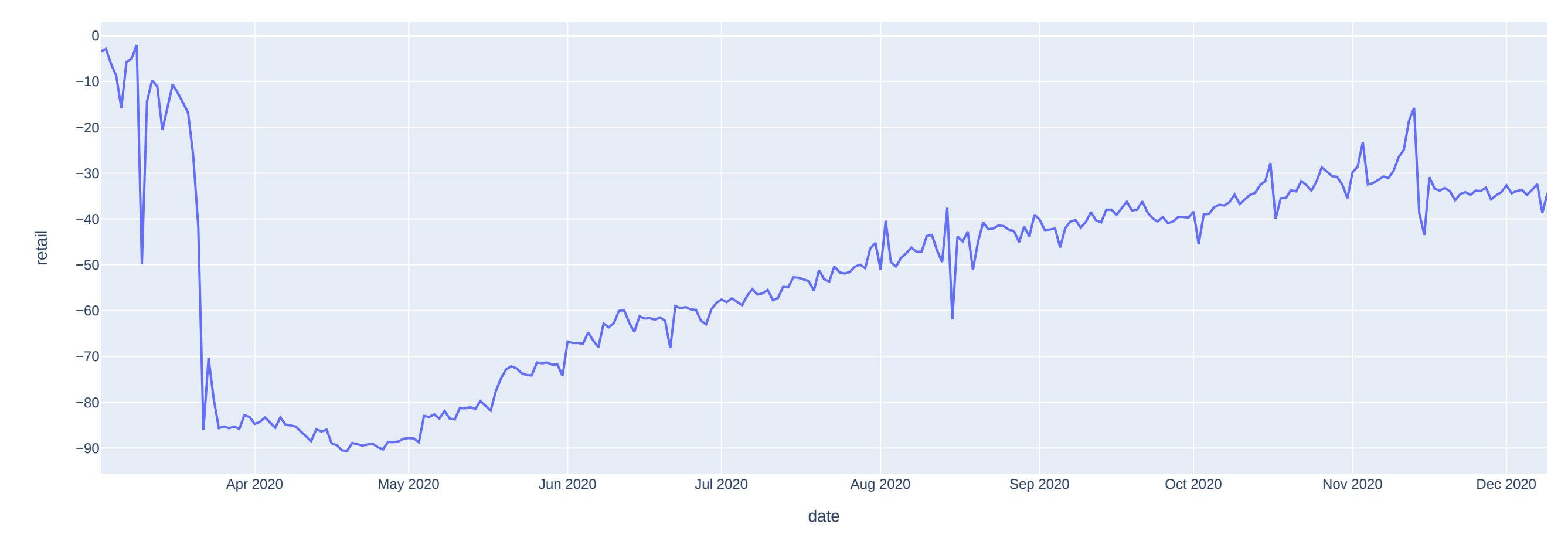
In [11]:

```python
import plotly.express as px
state_df = final_df.loc[final_df.state=="Delhi"].copy()
fig = px.line(state_df, x='date', y='cases')
fig.show()
```



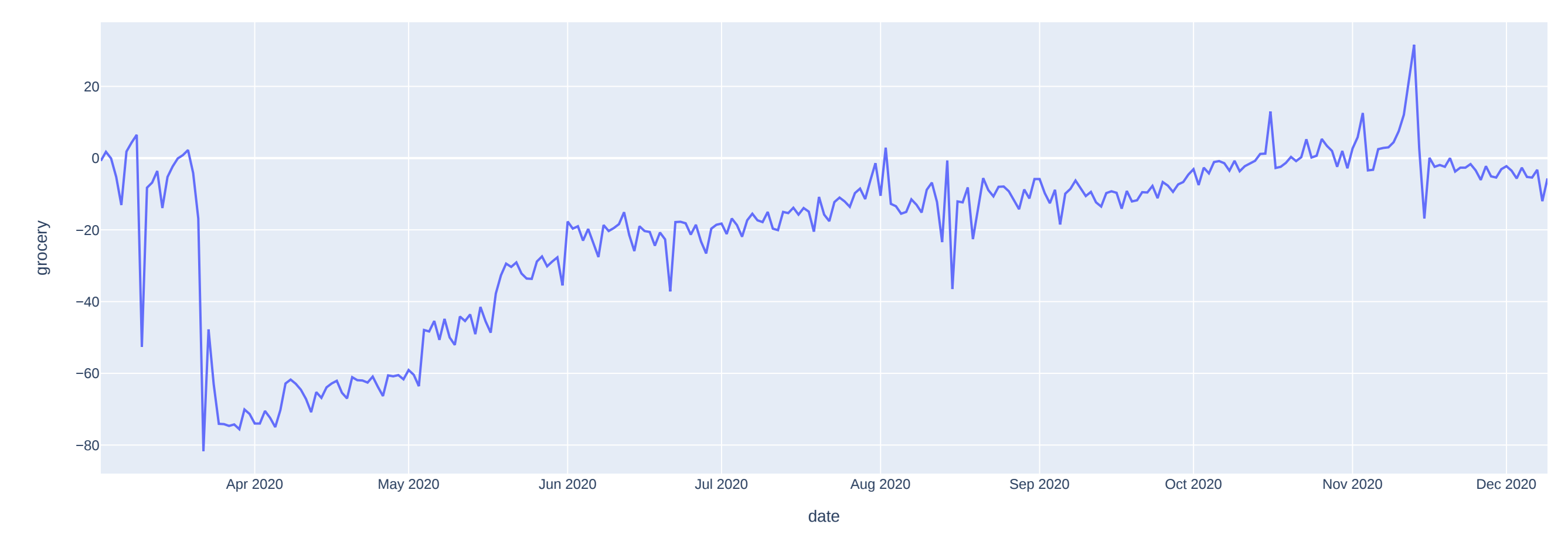
In [42]:

```python
fig = px.line(state_df, x='date', y='retail', labels = {'retail'})
fig.show()
```



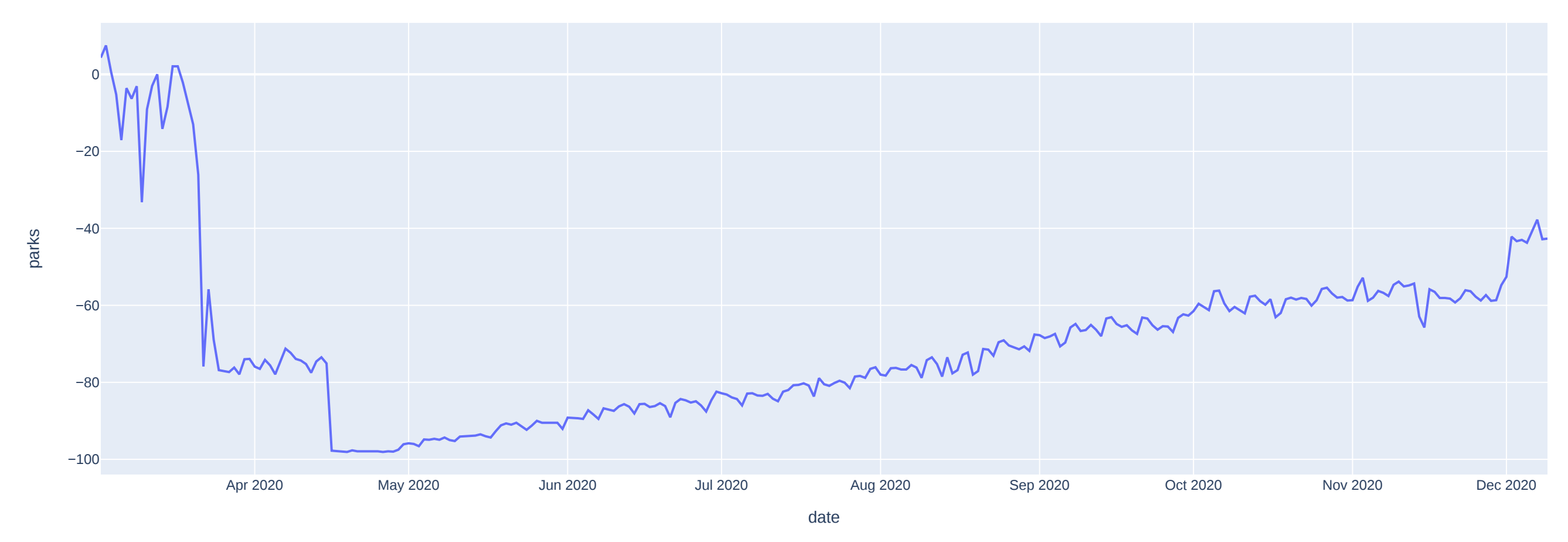
In [13]:

```python
fig = px.line(state_df, x='date', y='grocery')
fig.show()
```



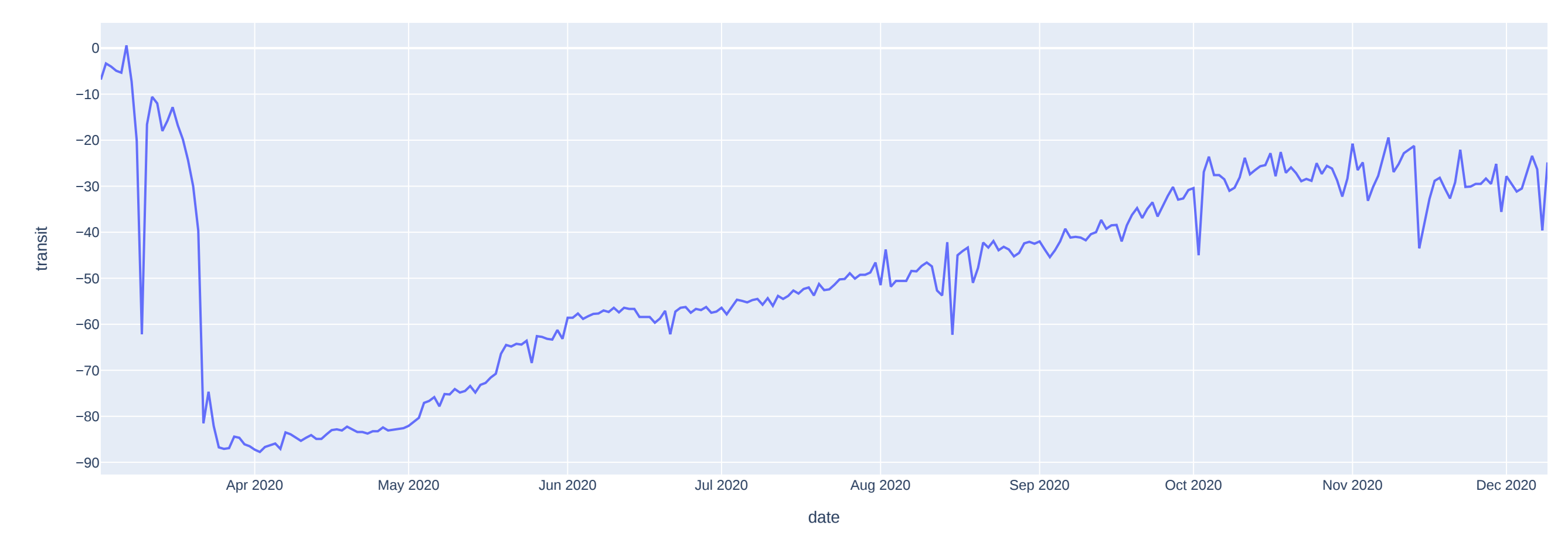
In [14]:

```python
fig = px.line(state_df, x='date', y='parks')
fig.show()
```



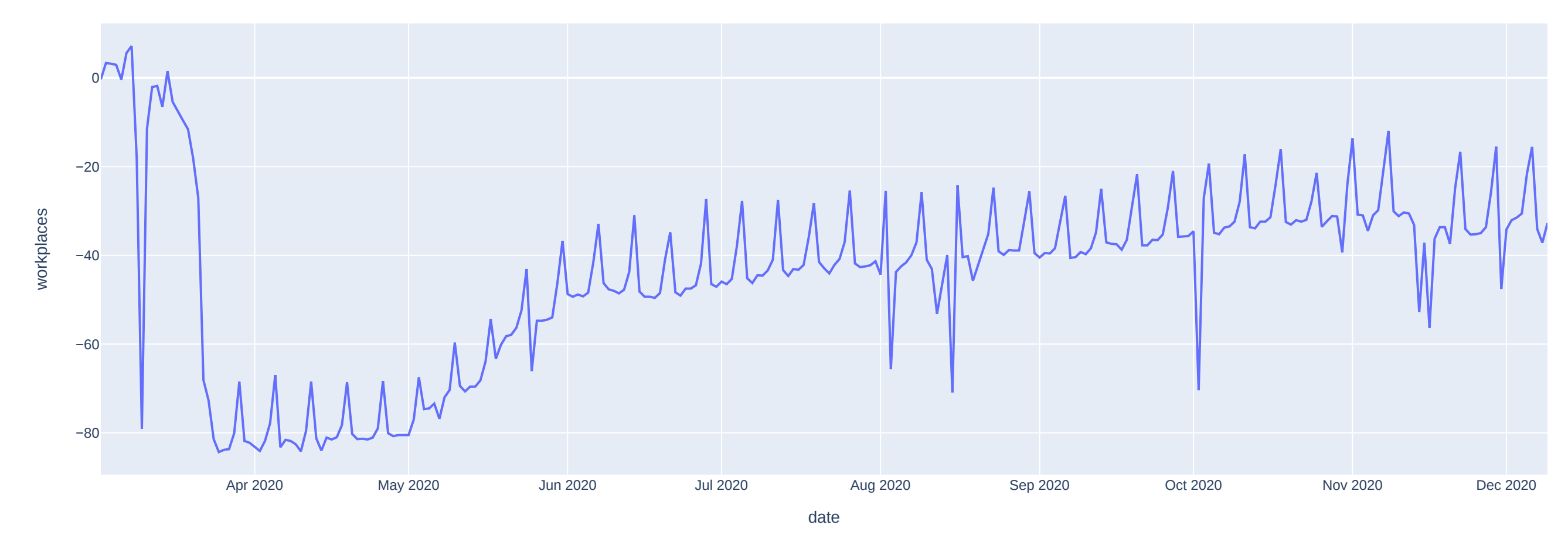
In [15]:

```python
fig = px.line(state_df, x='date', y='transit')
fig.show()
```



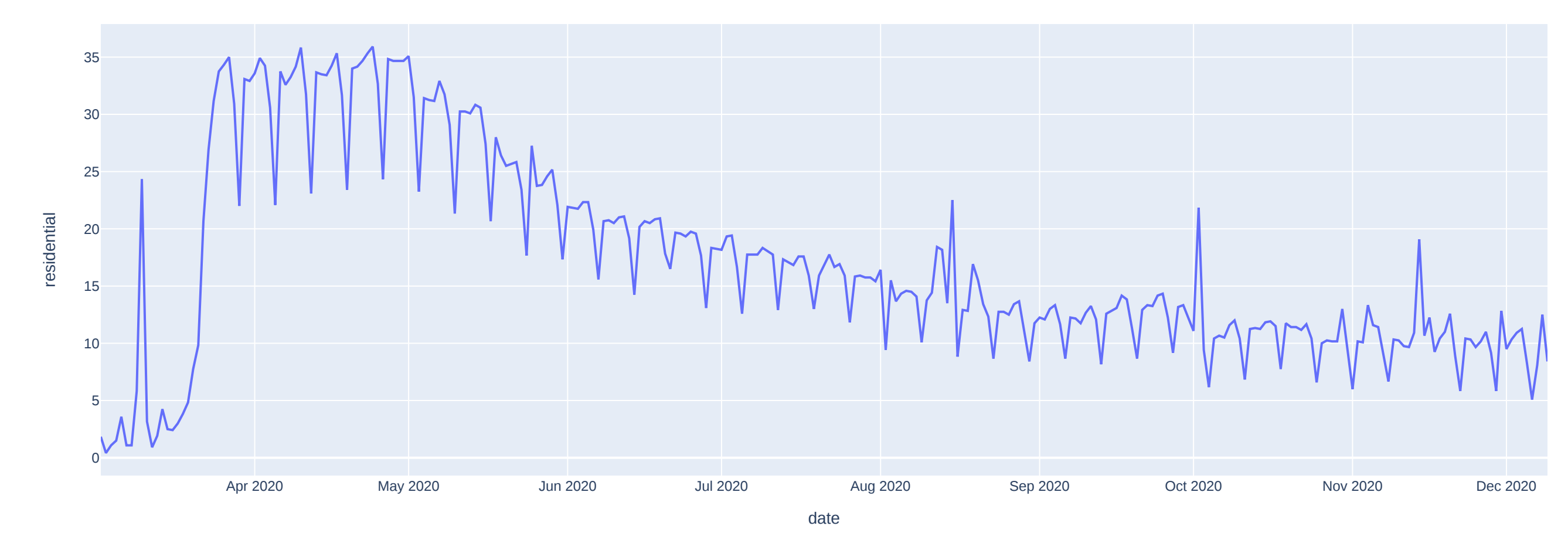
In [16]:

```python
fig = px.line(state_df, x='date', y='workplaces')
fig.show()
```



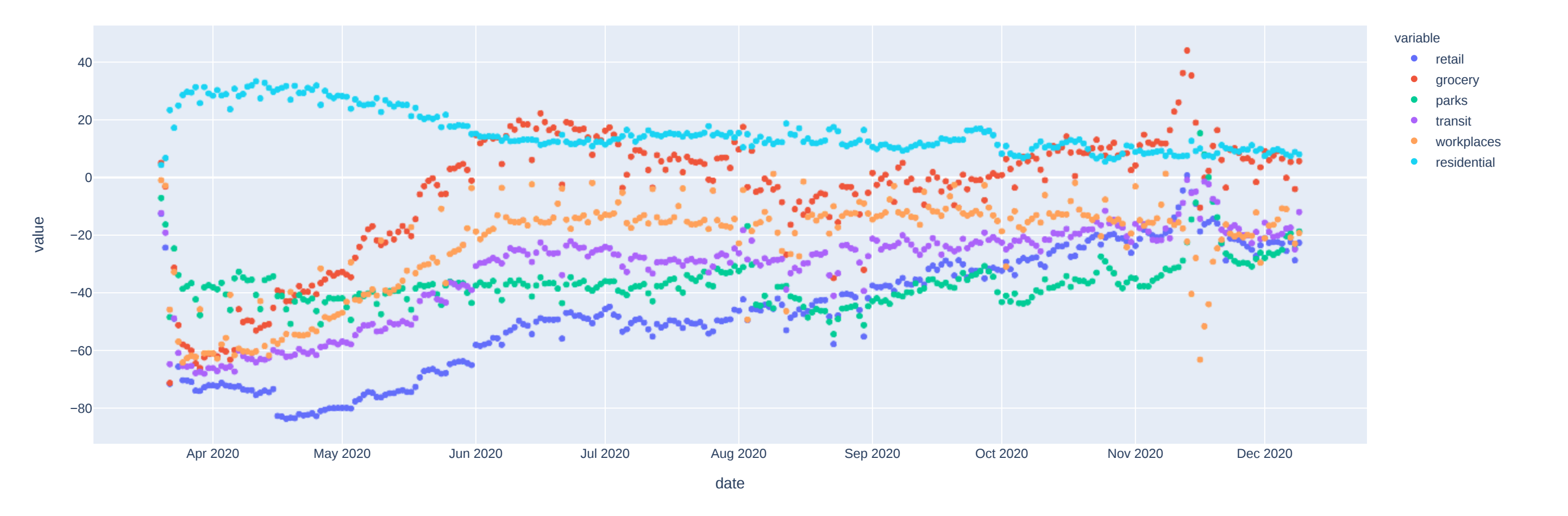
In [17]:

```python
fig = px.line(state_df, x='date', y='residential')
fig.show()
```



In [14]:

```python
fig = px.scatter(state_df, x="date", y=["retail", "grocery", "parks", "transit", "workplaces", "residential"])
fig.show()
```



In [15]:

```python
state_df.corr()
```

Out[15]:

	retail	grocery	parks	transit	workplaces	residential	cases	deaths
retail	1.000000	0.706770	0.516097	0.912014	0.612548	-0.892040	0.901732	0.922948
grocery	0.706770	1.000000	0.322583	0.898918	0.778340	-0.848115	0.512086	0.660717
parks	0.516097	0.322583	1.000000	0.466464	-0.042419	-0.358830	0.456837	0.338400
transit	0.912014	0.898918	0.466464	1.000000	0.774018	-0.958556	0.754678	0.871194
workplaces	0.612548	0.778340	-0.042419	0.774018	1.000000	-0.817758	0.432287	0.639895
residential	-0.892040	-0.848115	-0.358830	-0.958556	-0.817758	1.000000	-0.727108	-0.847125
cases	0.901732	0.512086	0.456837	0.754678	0.432287	-0.727108	1.000000	0.937159
deaths	0.922948	0.660717	0.338400	0.871194	0.639895	-0.847125	0.937159	1.000000

In [16]:

```python
state_df['new_cases'] = final_df['cases'].diff()
state_df['new_deaths'] = final_df['deaths'].diff()
```

In [17]:

```python
state_df.corr()
```

Out[17]:

	retail	grocery	parks	transit	workplaces	residential	cases	deaths	new_cases	new_deaths
retail	1.000000	0.706770	0.516097	0.912014	0.612548	-0.892040	0.901732	0.922948	0.020460	-0.145879
grocery	0.706770	1.000000	0.322583	0.898918	0.778340	-0.848115	0.512086	0.660717	0.061828	0.046584
parks	0.516097	0.322583	1.000000	0.466464	-0.042419	-0.358830	0.456837	0.338400	-0.174271	-0.242252
transit	0.912014	0.898918	0.466464	1.000000	0.774018	-0.958556	0.754678	0.871194	0.046085	-0.053282
workplaces	0.612548	0.778340	-0.042419	0.774018	1.000000	-0.817758	0.432287	0.639895	0.019733	-0.018426
residential	-0.892040	-0.848115	-0.358830	-0.958556	-0.817758	1.000000	-0.727108	-0.847125	-0.021987	0.076968
cases	0.901732	0.512086	0.456837	0.754678	0.432287	-0.727108	1.000000	0.937159	0.194203	0.005087
deaths	0.922948	0.660717	0.338400	0.871194	0.639895	-0.847125	0.937159	1.000000	0.236588	0.066140
new_cases	0.020460	0.061828	-0.174271	0.046085	0.019733	-0.021987	0.194203	0.236588	1.000000	0.963599
new_deaths	-0.145879	0.046584	-0.242252	-0.053282	-0.018426	0.076968	0.005087	0.066140	0.963599	1.000000

In [ ]:

In [ ]: